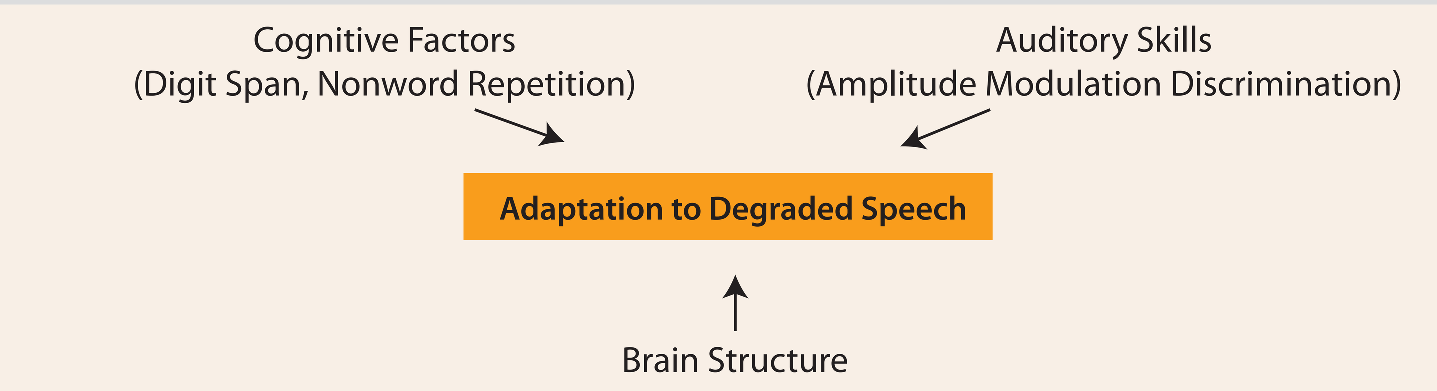
Auditory Non-speech Skills Predict Individual Differences in Adaptation to Degraded Speech

Julia Erb[1], Molly Henry[1], Frank Eisner[2] & Jonas Obleser[1]

[1]Auditory Cognition Research Group, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [2]Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands
jerb@cbs.mpg.de

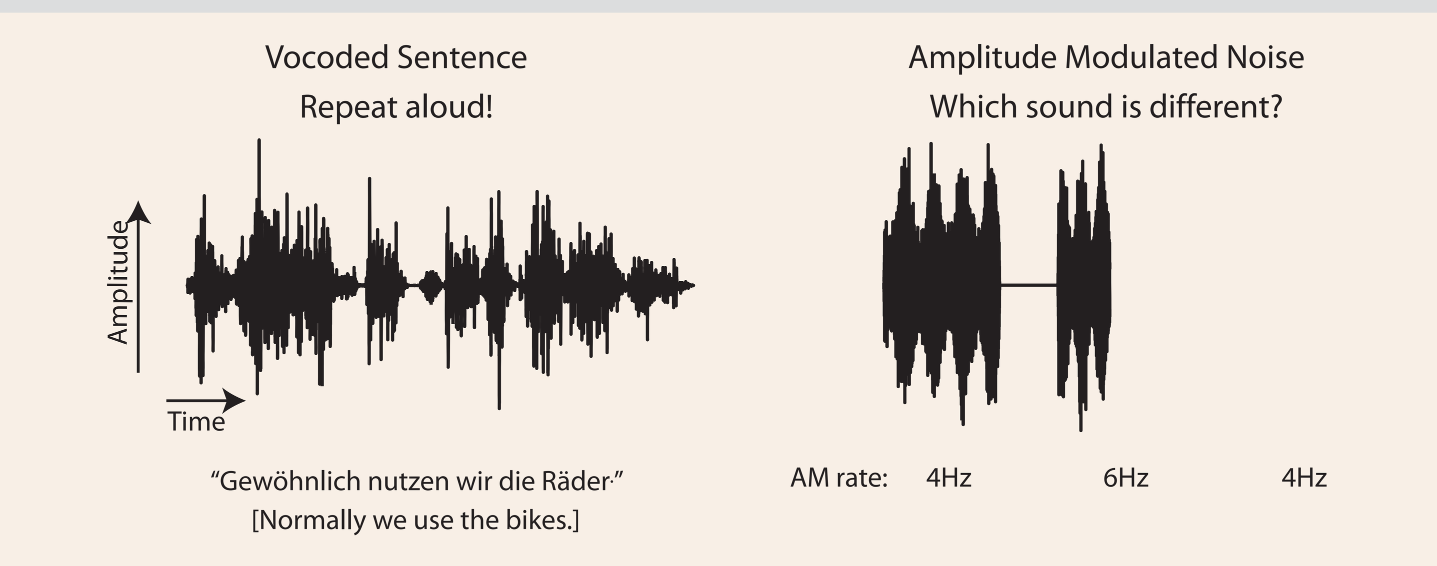
Introduction

- Noise-vocoded speech is a spectrally highly degraded signal but it preserves the temporal envelope of speech [1]. Noise-vocoding has been used to simulate cochlear-implant (CI) transduced speech in normal-hearing subjects.
- Listeners vary considerably in their ability to adapt to vocoded speech. Currently it is unclear what drives the adaptation to this degraded speech signal. Here, we hypothesized that individual differences in adaptation to vocoded speech could be predicted by cognitive, non-speech auditory and neuroanatomical factors.



Methods

- We tested 18 normal-hearing, right-handed adults (aged 22–30 years, 9 female) in a short-term vocoded speech learning paradigm. Stimuli were 100 low predictable German SPIN sentences [2], recorded by a female native speaker of German. Intelligibility was reduced using 4-band-noise-vocoding [3].
- Participants repeated after each sentence what they had understood. Vocoded speech perception scores (% correct) over time were fitted with a linear curve of which the slope was taken as measure of adaptation to vocoded speech (adaptation slope; see Figure 1B/C).
- Non-speech auditory skills were assessed using an amplitude modulation (AM) rate discrimination experiment. In a 3-Alternative forced choice paradigm, participants listened to amplitude modulated white noises; modulation rates were centered on the speech-relevant rate of 4Hz (range 2-6Hz, 8 linearly spaced levels, 160 trials total). Individual discrimination thresholds were calculated as just-noticable difference (JND) in Hz.



- Working memory capacities were evaluated using auditory forward and backward digit span and a nonword repetition test [4].
- Structural brain scans were collected on a 3T Trio TIM scanner. Data were analyzed using voxel-based morphometry running in SPM8 (segmentation into grey and white matter, creation of group-specific templates using DARTEL [5], spatial normalization to MNI-space, smoothing at 10mm FWHM). Multiple regression analysis was carried out with adaptation slope as covariate of interest and sex, age, time since acquisition and coil used for scanning as nuisance variables.

Results

1 Perceptual Adaptation to Vocoded Speech

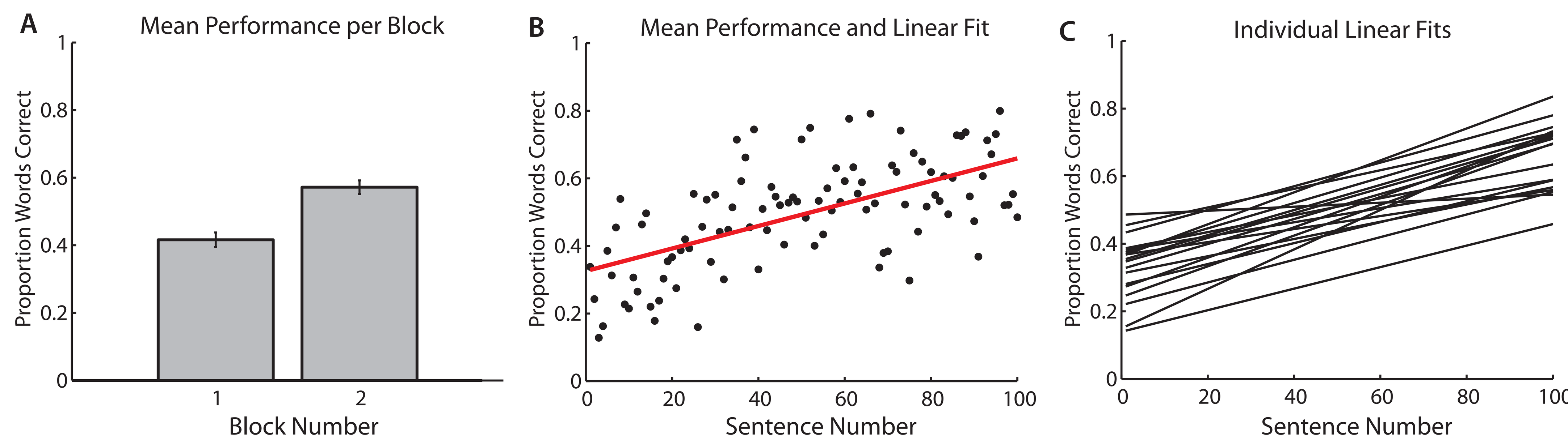


Figure 1: Perceptual adaptation to vocoded speech. (A) Average performance in the first and second half of the experiment (showing that learning took place), (B) linear fit to the average performance over all trials and (C) linear fits to individual speech perception scores, showing individual differences in adaptation to vocoded speech. The slope of the linear fit was taken as measure for perceptual adaptation to vocoded speech (adaptation slope).

2 Correlation between Adaptation Slope and AM Discrimination Thresholds

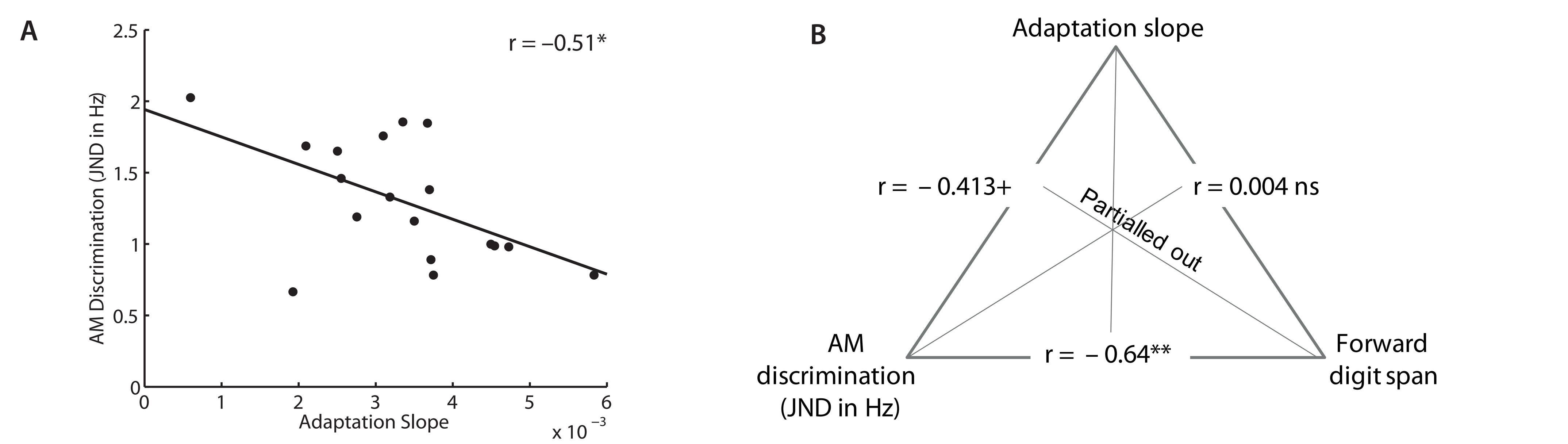
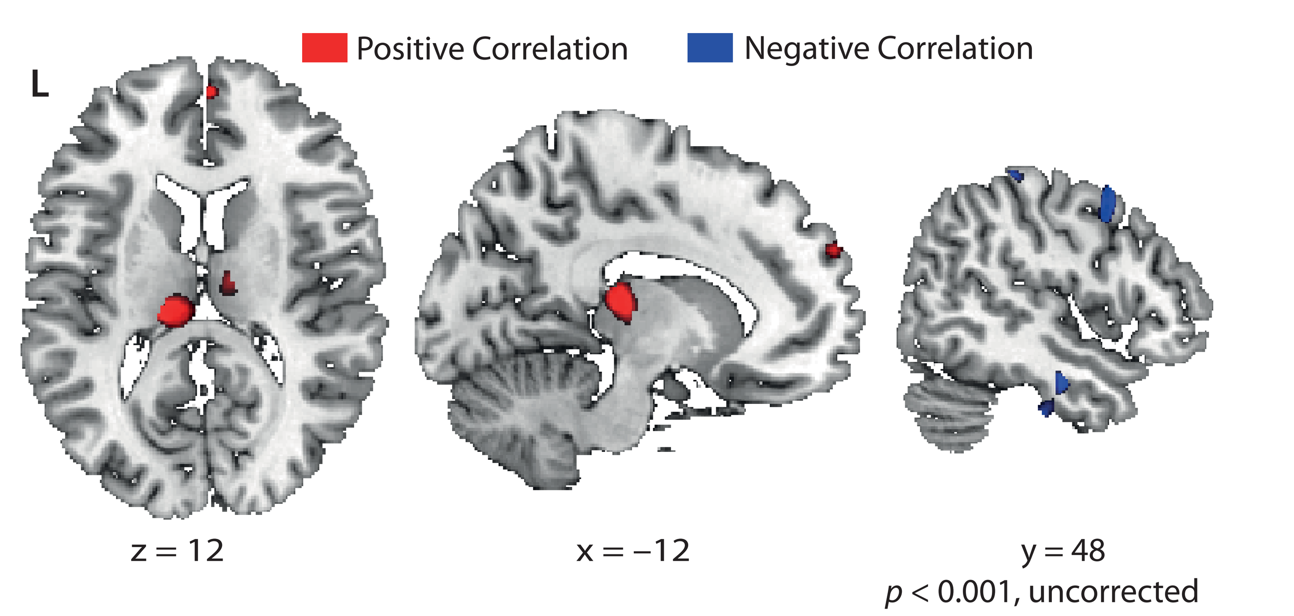


Figure 2: (A) Adaptation slope and AM discrimination threshold correlate significantly. No significant correlation was observed between adaptation slope and digit span or nonword repetition measures. (B) Partial correlation coefficients between adaptation slope, AM discrimination threshold and forward digit span (** $p < 0.01$, * $p < 0.05$ level, + $p < 0.1$, ns = non significant).

3 Brain Morphometry Predicts Adaptation to Vocoded Speech



Location	Z-Score	Extent (mm³)
Adaptation to vocoded speech (slope)		
Positive correlation		
L thalamus (pulvinar)	3.56	280
R thalamus ‡	3.19	17
R frontal pole ‡	3.53	27
Negative correlation		
R middle frontal gyrus	3.92	152
R inferior temporal gyrus	3.68	66
R postcentral gyrus	3.51	60

‡ did not meet cluster criterion

Figure 3: Results of voxel-based morphometry analysis. Adaptation slope correlates positively with regional volume in left thalamus (pulvinar) and negatively with right-lateralized premotor and motor areas as well as inferior temporal gyrus.

4 Jack-knife reanalysis of VBM results

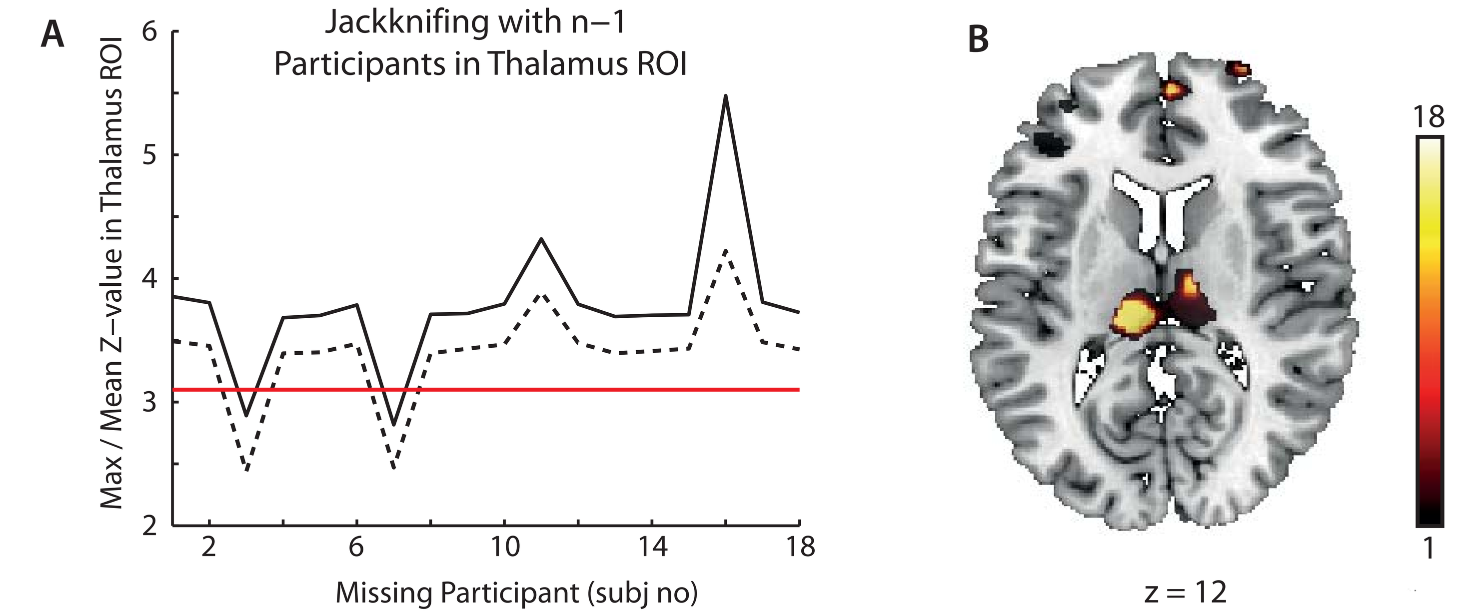


Figure 4: Iterative (jack-knife) reanalysis of VBM multiple regression analysis to test reliability. (A) Data points indicate maximum Z-values (or mean Z-values as dotted line) within region of interest (ROI) in the left thalamus when one subject left out at a time (n = 17). ROI in the left thalamus was defined based on the primary analysis including all 18 subjects. The threshold of Z = 3.1 is depicted in red, showing that Z-values are above threshold for 16 out of 18 subsamples (each one including n-1 participants) which can be considered a reliable result. (B) Colors code for the number of subsamples (each one including n-1 participants) that have a threshold of Z > 3.1 when conducting a whole-brain iterative reanalysis.

Conclusions

- Our results corroborate that fast perceptual learning of vocoded speech (exposure to only 100 sentences) is achieved in the absence of feedback and that there is considerable individual variability [6,7].
- Individual sensitivity to AM rate is predictive of vocoded speech learning. This suggests that adaptation to vocoded speech benefits from AM rate discrimination skills as both listening situations require the use of envelope cues in the auditory signal.
- AM rate discrimnation skills are a more robust predictor of adaptation to vocoded speech than digit span and nonword repetition measures.
- The ability to adjust to degraded speech is reflected anatomically in an increased volume in the left pulvinar which is strongly connected to the auditory and prefrontal cortex [8].
- It will be important to verify these results in cochlear implant patients. Structural brain scans as well as measures of auditory skills could be used to distinguish prospectively good from poor cochlear implant users and to shape individual strategies accordingly.

References

[1] Shannon et al. (1995). Science 270, 303-304.
[2] Kalikow et al. (1977). J Acoust Soc Am 61, 1337-1351.
[3] Rosen et al. (1999). J Acoust Soc Am 106, 3629-3636.
[4] Mottier et al. (1951). Folia phoniatrica 3, 170-177.
[5] Ashburner et al. (2000). Neuroimage 11, 805-821.
[6] Davis et al. (2005). J Exp Psychol Gen 134, 222-241.
[7] Eisner et al. (2010). J Neurosci 30, 7179-7186.
[8] Behrens et al. (2003). Nat. Neuroscience 6, 750-757.